DOCUMENT RESUME

ED 144 502 HE 009 290

TITLE Evaluating Innovation. An Occasional Report on Innovation.
INSTITUTION California State Univ. and Colleges, Los Angeles. Office of the Chancellor.
PUB DATE [77]
NOTE 5p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Educational Assessment; *Evaluation Methods; *Evaluation Needs; *Higher Education; Innovation; Personnel; *Program Evaluation; Relevance (Education); *Research
IDENTIFIERS *California State University and Colleges

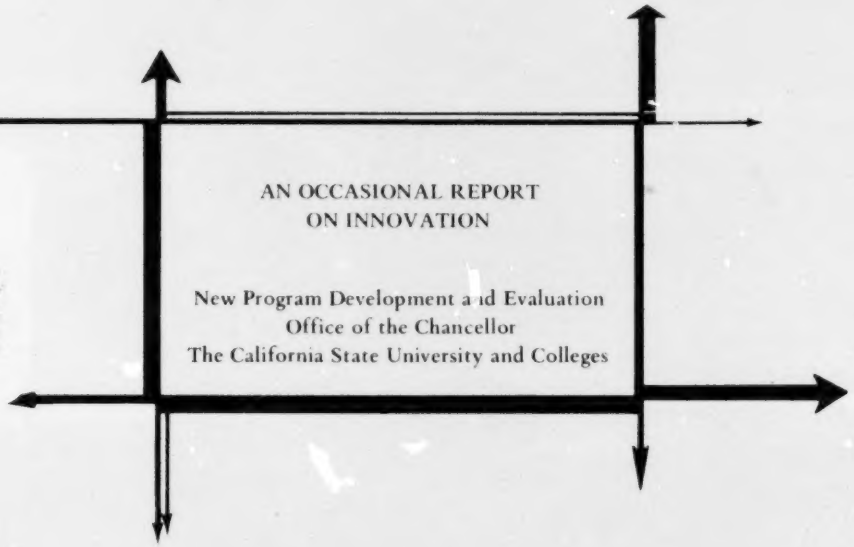
ABSTRACT

Evaluation is a process of systematically gathering information on a regular, continuing basis to make informed judgments about the merits of a program or an activity. This discussion of evaluation issues and procedures represents an attempt to place the role of evaluation in perspective and to provide guidelines and suggest approaches, particularly within the context of projects sponsored by the California State University and Colleges Program for Innovation. The purposes of evaluation are discussed, including the determination of a program's success in attaining its original objectives, as well as identifying unanticipated consequences. Several dimensions of evaluation are outlined dealing with distinguishing between types of evaluation; determining the focus of the evaluation; and the methodology involved. The necessity for a planned evaluation design is discussed to ensure that the program contains the necessary elements upon which the evaluation will be based. The makeup of the project staff is outlined, the chief guideline being the scope and format of the project should determine who conducts the evaluation.

**********
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
**********

ED144502

# Evaluating Innovation


AN OCCASIONAL REPORT
ON INNOVATION
New Program Development and Evaluation
Office of the Chancellor
The California State University and Colleges

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

Evaluation is a process of systematically gathering information on a regular, continuing basis in order to make informed judgments about the merits of a program or an activity. To some innovators, however, evaluation is a threatening thing and is better out of sight, out of mind. To others, even questionable evaluative data is considered "divinely inspired" and a preoccupation with end-of-program findings leads to a disregard for program processes. With reasonable approaches to data collection and interpretation, both of these extremes can easily be avoided.

The following discussion of evaluation issues and procedures represents an attempt to place the role of evaluation in perspective and to provide guidelines and suggest approaches, particularly within the context of projects sponsored by The California State University and Colleges Program for Innovation. Persons interested in pursuing new approaches to teaching and learning, whether through a formal project or not, may find the discussion beneficial.

HE009290

## Why Evaluate?

The purpose of evaluation is not simply to satisfy demands for accountability imposed by legislatures, university departments or a central office. Basically, a program is evaluated in order to determine its success in attaining its original objectives as well as to identify unanticipated consequences. Impact of the program can be determined in terms of student performance, student retention, the student's view toward a discipline, and other similar concerns.

Certain questions are basic: If the effort was worthwhile, should others be encouraged to implement similar activities? Were all aspects of the program worthwhile, or should some elements be revised? A project may be efficiently managed and give the appearance of success but meet only minor objectives with inconsequential results or results irrelevant to the student. Too often, projects employing efficient means attain carelessly examined ends.

Important to any evaluation are the overriding goals of the organization or program within which the innovation is being undertaken. In the case of the CSUC Program for Innovation, some of the broadly defined goals of this program are:

- Exploring ways to increase student learning
- Improving faculty's satisfaction with teaching
- Increasing student learning options
- Facilitating constructive institutional change and renewal
- Providing a systemwide approach to cumulative, constructive change and renewal

- Identifying the costs and benefits of change and renewal
- Encouraging policies and procedures which will encourage innovation and change

The goals of individual projects should interface with those of the organization within which the project occurs, though certain goals are obviously emphasized within a given type of project.

## Dimensions of Evaluation

Evaluation may be viewed through three different dimensions, each of which involves the collection of different types of data. The first dimension distinguishes between three types of evaluation – formative, summative and impact evaluation. Formative evaluation involves gaining feedback for program revisions and modifications, including adjustment of program objectives during the developmental or formative stages of the project. Particular attention is given to collecting data regarding the target population, performances at each state of the program, the conditions under which these performances vary and the appropriateness of the program objectives within the context of the program's evolution.

Summative evaluation is directed towards validation of the project upon completion, and usually involves the use of comparative data concerning the overall merits of the particular project in relation to similar projects for similar audiences. Impact evaluation involves a consideration of the project and its results within the context of the department, campus or system. In other words, impact evaluation is concerned with the overall consequences of the project whether or not the program achieved its objectives.

The second dimension concerns the focus of the evaluation: the process of the activity or its project. A good evaluation, however, should be concerned with both process *and* product. A program cannot be considered worthwhile if it achieves its objectives in a manner that may be destructive to the subjects – for instance, where modest learning gains are achieved at the cost of creating antipathy toward the discipline subject matter.

The third dimension of evaluation deals with methodology. In the grossest terms, this concerns the interplay between objective and subjective data. Again, this is not an "either-or" proposition in practice. The objective collection of data and the subjective assessment of that which is non-quantifiable, such as perceptions and feelings, are both important, but must be clearly differentiated. By a process of triangulation, the same subject matter can be viewed from the different perspectives while employing different data collection techniques.

## Evaluation Design

A carefully planned evaluation design is imperative to the implementation of a systematic evaluation strategy. Moreover, the evaluation plan must be considered in the *program design* to ensure that the program contains the necessary elements upon which the evaluation will be based. Well-defined objectives thus lie at the heart of both the program and evaluation designs.

Objectives may be stratified in terms of the following program implications:

- Impact on faculty
- Impact on student
- Impact on institution (and system)
- Impact on curriculum

The evaluation design then may be built upon the objectives of the program itself. The evaluation and the program design should consider the consequences of stressing one objective over another and be flexible enough to discard objectives which are determined to be unattainable or inappropriate to the target audience. Similarly, the project and its evaluation should be prepared to develop *new* or additional objectives where appropriate as the program evolves.

The actual evaluation plan or strategy will vary according to the type of program involved. One type of project, such as the comparison of two only slightly different learning strategies, may require a more rigorous program and evaluation design than another, for example, which might depend more upon student and faculty interviews to assess attitudes toward a cluster college program.

Some projects require a strict experimental design to provide appropriate evaluation data; others may be evaluated in a less rigorous fashion. Stanley and Campbell's (1966) excellent description of experimental and quasi-experimental designs and Popham (1975) are useful reading in this area.

Project developers can find a number of different people on campuses who can help them with their evaluation plans. Test officers and faculty in schools of education and departments of psychology and sociology often can provide assistance with test instruments. CSUC campus coordinators for innovation can assist in identifying individuals with specific expertise. Graduate and undergraduate students may be available to perform evaluation related activities as a portion of their academic program. Other project directors also can serve as advisors.

One of the purposes of the evaluation component of a project is to provide a historical perspective on what occurred, and why. Projects evolve within different contexts stimulated by pressures of the moment, and thus take different directions and approaches. Unless a conscious effort to assess the project at key points in the process is built into the evaluation design, the reasons for shifting project directions become lost, as well as the record of the process of adjustment to the new direction. The reasons why a project changed during its development may be just as important to others as the outcomes achieved.

In essence, the purpose of an evaluation is to find out if you are doing what you think you are doing during the course of the project (formative); and, at the end of the project, to tell you whether you actually did what you thought you did (summative). It should also tell you whether the whole effort or only a portion of it was worthwhile (impact). The project may have proven to be technically successful, but of little impact. The benefits achieved may not be sufficient, given the human and dollar costs invested. If the project is deemed worthwhile, the evaluation should pay attention to ways in which others might attempt to implement it and under which circumstances.

Determining how and when to collect what kinds of data is of primary importance to the evaluation of a program. Some basic questions need to be asked in considering what data to collect. The purpose of the data, from whom the data needs to be collected, and how much data to collect are all important questions that must be considered when designing the evaluation plan.

Avoid the "survey research syndrome" which tends to lead to the collection of data for the sake of collection, whether relevant or not. A policy of lean data collection is always preferable provided all essential data are collected. Distinguishing between the types of data related to performance (or process) and consequence (or outcome) is useful in considering data collection plans. The two data collection efforts must be mutually supportive. One cannot simply collect data on student performance and expect to learn whether or not an objective related to changing faculty roles has been met.

It may be useful to think of data needs in two categories. For example, one data set can relate to *learning outcomes* and is student focused. Here diagnostic testing, pre- and post-testing, grades and other assessments, and observation of a control group may be required. Retention rates and assessment of affective behavior such as attitude toward the subject matter, as compared to subject matter learning, relate to learning outcomes in a broader sense.

Other kinds of data may be grouped under the rubric of the *impact* of the project. Here the human, fiscal and physical costs of the project may be considered. Data can be collected on the use of faculty time in the innovation, curriculum change caused by the innovation, and procedures and policies which require modification for the innovation to be continued and extended. Most important in assessing project impact is the collection of comparison or base-line data. If judgments are to be made about the dollar costs of instruction in the new method, costs in the traditional instructional mode should be known. How faculty time is used in the traditional mode also should be known before judgments about the use of time in a new method can be made. Unfortunately, very few comparative analyses of traditional modes of instruction and student attitudes toward them, or of the uses of student and faculty time in a traditional program have been undertaken, providing a special problem in evaluating new approaches.

It is also important to ask the right question at the right time. Pre-tests at the outset of an experimental instructional effort are often neglected, and then found to be needed when it is too late to collect this important base-line information. Decisions concerning the need for pre-program data must be made at the outset of program planning and included where appropriate in the evaluation design.

The type of assessment instruments which are most appropriate for both the program and the evaluation design must be considered at the outset. Too often educators feel that they must develop new instruments for the evaluation when standardized instruments may contain appropriate items of tested validity and reliability, while at the same time providing comparative data. The design of "home-grown" instruments must always be balanced against the legitimate use of those which are already extant and validated, carefully scrutinizing their appropriateness to the particular program objectives being assessed.

Another important consideration pertains to sample size and frequency of testing. Students and faculty have become increasingly resistant to the "measurement fever" that has spread throughout higher education. Some of this resistance is deserved. Educators too often inundate students and each other with needless surveys. Likewise, whole populations are sometimes assessed when smaller samples would be sufficient. Consideration must be given to both over- and under-sampling in designing useful and efficient evaluation plans.

An alternative to the use of questionnaires or other instruments which may result in sample resistance is the use of unobtrusive measures which do not involve contacting students or faculty directly. Drop-out rates, attendance records, change-of-major patterns, and rates of material usage fall into this category of data collection. Behavioral indices may also be obtained in this manner. Student observation, simulated activities, and the mixing of students with one another and monitoring their topics of conversation are all appropriate in this regard. For a discussion of unobtrusive measures, see Webb *et al.*, (1966).

## Who Evaluates?

An important issue concerns who should conduct the evaluation. At times, it may seem that the only possible definitive evaluation can come from the Big Evaluator in the Sky. This may be ultimately true, but, practically speaking, there are a number of different potentially effective evaluators. Students, both within and outside a project, are a good source of evaluation for some aspects; faculty colleagues are another source; a panel of consumers of the product (usually students who have completed the program) could also be used effectively. The scope and format of the project should determine who conducts the evaluation.

Most project staffs in the CSUC Program have collected data themselves and often have performed the entire

evaluation function internally. Some have combined internal with external evaluation. A mix of evaluations – students, colleagues and an independent third party evaluator – is most desirable so that different perspectives come to bear, though ultimately the preparation of a consolidated evaluation report should be concentrated in one or two individuals.

External evaluators, when used, should be brought into the project at the earliest possible point in time. Though they need not create the evaluation design, they should have a thorough understanding of it and be able to suggest improvements. They also should have a clear understanding of the project's objectives and collaborate in the definition of criteria upon which the meeting of those objectives is to be determined, for they must report whether or not the objectives are actually being met.

The outcome of the evaluation should be addressed to the implications of the project, not only within its own context and objectives, but also within the larger world around it. One should be able to determine what quality has been achieved, and can likely be achieved in the future, and at what costs. For most innovations involving a change of instructional strategies, the evaluation should address the question of whether or not a campus can carry on the program without special assistance. Related to this matter is a consideration of what kinds of reallocations of resources and efforts may be required to carry on the activity if it is judged to be worthwhile.

The implications of negative results should be considered early in the program. Often there is value in failure, particularly when it is known *why* something failed. An idea may simply have been ahead of its time in a particular setting or only certain aspects of a program may have failed. Failures and the reasons for them should be clearly delineated and considered in the writing of recommendations and implications.

Distinctions can be made between the implications an innovative effort has for the institution as opposed to the personal value gained by its designers. Too often, formal project reports ignore what may be the most significant outcome of a project – the experience of the faculty involved. A project may have required faculty members to become familiar with different disciplines or instructional methodologies. Additionally, new ways may have been learned in which to work with students, give them counsel, or assess their competencies.

## Dissemination

The dissemination responsibility of those who support projects and programs, as well as the participants themselves, is often forgotten. Too often a single copy of a project's final report is filed on a library shelf signifying the end to the responsibility for the project director and the sponsor. A good evaluation plan should have, as a major component, an information feedback mechanism for those interested at the departmental, institutional and system level. Where appropriate, the higher education community as a whole is the audience. Clear, concise reporting, in a form suitable for general publication, is perhaps the most effective single ingredient in dissemination. Workshops, presentations to professional associations in the discipline, and word of mouth are important as well. But it must be remembered that dissemination begins at home – colleagues in the department, on the campus and in the system should be aware of what is being attempted and what has resulted. The hall corridor can be just as great a communications abyss as a continent.

Evaluation should not be perceived as threatening or negative. Its purpose is the improvement of the project and its methodology should be rooted in common sense. The evaluation plan is a sensing device to tell when things are going well and when they are not. But, throughout the activity or program, don't lose sight of the reason for the effort in the first place. It is a much easier task to create a good evaluation design than to develop a sound and worthwhile program.

## References


Borich, G. D. (Ed.). *Evaluating Educational Programs and Products.* Englewood Cliffs, New Jersey: Prentice-Hall 1974.

Campbell, D. J. and Stanley, J. D. *Experimental and Quasi Experimental Designs for Research.* Chicago: Rand McNalley, 1966.

Popham, W. J. "Program Fair Evaluation – Summative Assessment of Instructional Sequences with Dissimilar Objectives," *NSPI Journal,* 1969, 8 (6), 6-9.

Popham, W. J. (Ed.). *Evaluation in Education.* Berkeley, California: McCutchan Pub. Corporation, 1974.

Popham, W. J. *Educational Evaluation.* Englewood Cliffs: Prentice-Hall, 1975.

Webb, E. J., Campbell, D. T., Schwartz, R. D., and Sechrest, Q. *Unobtrusive Measures: Nonreactive Research in the Social Sciences.* Chicago: Rand McNally, 1966.

Weiss, C. H. *Evaluating Action Programs: Readings in Social Action and Education.* Boston: Allyn and Bacon, Inc., 1974.


★ ★ ★ ★


This short occasional paper is the result of a wide-ranging discussion. Participating were:

*Harold Best, CSU, Fresno*
*Jack Guthrie, CSU, Hayward*
*Robert Picker, San Francisco State University*

And from the Office of the Chancellor, Division of New Program Development and Evaluation:

*Toni A. Howard*
*David E. Leveille*
*Glenn F. Nyre*
*Clare Rose*
*John M. Smart, recorder*